ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертационную работу Русанова Михаила Андреевича

«Интеллектуальные методы обнаружения несанкционированных вторжений в операционные системы»,

представленную к защите на соискание ученой степени кандидата технических наук по специальности 2.3.5 — «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей»

1. Актуальность

Диссертационная работа посвящена обнаружению несанкционированных вторжений по системным событиям операционных систем семейства Linux и оценке риска компрометации учетных данных. Под системными событиями понимается запуск и завершение процессов, доступ к файловой системе, изменение прав и привилегий, действия пользователей и сервисных учетных записей и т.д. Актуальность такой постановки определяется тем, что в современных инфраструктурах значительная часть действий пользователя, приложения или нарушителя отражается в журналах операционной системы и прикладных сервисов. При этом события поступают непрерывно, а эффективность обнаружения атаки напрямую связана со временем реакции: выявление вредоносного действия после завершения атаки существенно менее ценно, чем его обнаружение и принятие соответствующих мер на ранней стадии.

С практической точки зрения существенным является выбранное в работе ограничение по вычислительным ресурсам. Многие современные методы анализа журналов предполагают использование значительных ресурсов – вычислений на GPU, использования сложной распределенной инфраструктуры и т.д. В диссертации, напротив, исследуется возможность выполнять первичную классификацию событий на CPU, используя сравнительно компактные нейросетевые модели и небольшое текстовое окно. Такая постановка представляется обоснованной: в корпоративной системе мониторинга полезен недорогой ранний фильтр, способный быстро отделять основную массу штатных событий от событий, требующих более детального анализа. Он не заменяет корреляцию событий и расследование инцидента в SIEM, но может уменьшить объем последующей обработки и сократить задержку обнаружения.

Отдельное направление работы связано с анализом паролей. Здесь задача сформулирована как выявление потенциально компрометируемого пароля не только по формальному соответствию парольной политике, но и по статистическим характеристикам его структуры. Такая задача также имеет практическое значение, поскольку формально сложный пароль может оставаться предсказуемым и быть близким по структуре к ранее скомпрометированным паролям.

2. Краткое содержание работы

Объём диссертации составляет 130 стр., включая 30 рисунков и 20 таблиц. Текст состоит из введения, трёх глав, заключения и списка литературы (108 источников). Во ***введении*** сформулированы цель и задачи исследования, объект и предмет работы, научная новизна, теоретическая и практическая значимость, положения, выносимые на защиту, сведения о публикациях, апробации и внедрении.

В ***первой главе*** выполнен обзор методов обнаружения вторжений по системным журналам, рассмотрены ограничения ресурсоемких нейросетевых и графовых подходов, обсуждаются тактики атак в терминах Cyber Kill Chain и MITRE ATT&CK, а также задача оценки надежности пользовательских паролей.

Во ***второй главе*** предложена двухступенчатая схема анализа системных событий. Сначала выполняются очистка, нормализация и токенизация записи журнала, после чего бинарный классификатор определяет, является ли событие нормальным или вредоносным. Только события, признанные потенциально вредоносными, передаются второму, многометочному классификатору (модель семейства BERT). Полноразмерная BERT-модель применяется как исходная, а компактная TinyBERT — как модель-ученик; для уменьшения вычислительной стоимости используются дистилляция знаний и квантизация. В эксплуатационных экспериментах максимальная длина входа ограничивается 128 токенами, то есть фактически рассматривается небольшое локальное окно внутри отдельного события, а не длительная временная история поведения узла.

В этой же главе представлен метод оценки пароля. Пароль представляется вектором признаков, включающим длину, количество используемых групп символов, долю уникальных символов, долю цифр и специальных символов, среднее расстояние между соседними символами на клавиатуре и частоту смены алфавитов. Для классификации исследованы Decision Tree, Random Forest, Naive Bayes и XGBoost; в итоговой реализации выбран Decision Tree.

Третья глава посвящена программной реализации и экспериментальной оценке представленных в предыдущей главе результатов. Программный комплекс для анализа событий был реализован на языке Python с использованием библиотек PyTorch и HuggingFace Transformers; для отдельных экспериментов также использовались библиотеки scikit-learn, pandas, numpy. Программный комплекс включает модули предобработки, обучения бинарного и многометочного классификаторов, дистилляции, квантизации, расчета метрик и потокового вывода. Для повышения производительности использована пакетная обработка событий, динамическое выравнивание последовательностей и группировка событий по длине. Парольный модуль интегрирован в стандартную подсистему Linux PAM посредством pam_exec.so: новый пароль передается локальному процессу анализа, после чего результат классификации влияет на разрешение или запрет его установки. Для комплекса анализа событий была предложена схема подключения к SIEM через дополнительный коннектор.

Также в этой главе представлены результаты экспериментов. Для первой задачи (классификация вредоносных событий) был построен синтетический набор данных на основе тестовой среды, эмулирующей реальную: после очистки получилось 114 851 событий, из которых 91 804 являются нормальными, а 23 046 – вредоносными. Использовались метрики precision, recall, F1, accuracy, ROC-AUC, PR-AUC, матрица ошибок и время выполнения inference. Бинарная классификация дала accuracy 0,99%, и такое же качество демонстрирует весь подход в целом. Для исследования результата по анализу паролей была создана сбалансированная выборка из 770 тыс. паролей, при этом «потенциально компрометируемые» были получены из известной утечки rockyou, а надежные были искусственно сгенерированы известными сервисами Password Generator Plus, Bulk Password Generator и Check Your Password (Kaspersky). Получено значение accuracy составило 97,87 %.

В ***заключении*** ещё раз приводятся основные результаты работы.

3. Перечень научных результатов

Основные научные результаты соответствуют задачам и обладают научной новизной. Результаты можно сформулировать следующим образом.

1. Метод обнаружения и классификации вредоносной активности в потоковых системных событиях Linux, включая нормализацию и семантическое упрощение, бинарную классификацию «нормальное/вредоносное» и многометочную классификацию по тактикам MITRE ATT&CK. В качестве нейросетевых моделей использованы BERT и TinyBERT; для получения компактной модели применяются дистилляция знаний и 8-битная квантизация. Метод ориентирован на анализ отдельных событий с небольшим размером входного окна и на работу в условиях ограниченных ресурсов.
2. Метод параллельной потоковой обработки системных событий без изменения архитектуры используемой нейросетевой модели, который осуществляет адаптивное формирование вычислительных пакетов, динамическое дополнение последовательностей до максимальной длины и предварительную группировку событий по длине токенизированного представления. В проведенных экспериментах совместное применение этих механизмов увеличило пропускную способность примерно в 2,23 раза относительно базовой схемы обработки.
3. Метод оценки вероятности компрометации аутентификационных данных операционных систем семейства Linux. Пароль представляется набором статистических и структурных признаков случайности, после чего выполняется классификация методами машинного обучения (сравниваются Decision Tree, Random Forest, Naive Bayes и XGBoost, выбран Decision Tree).

4. Достоверность и обоснованность результатов

Достоверность результатов обеспечивается экспериментами, сравнением нескольких моделей и использованием групп метрик.

Публикационная апробация соответствует уровню кандидатской диссертации. По теме работы указано 6 печатных работ; три публикации выполнены в изданиях, рекомендованных ВАК РФ, две работы опубликованы в международном рецензируемом журнале Big Data and Cognitive Computing и индексируются в WoS/Scopus. Наличие публикаций как в российских профильных изданиях, так и в международном индексируемом журнале является существенным положительным фактором диссертационной работы.

Кроме того, было получено два свидетельства о государственной регистрации программ для ЭВМ и акт внедрения из АО «Газпромбанк». Для кандидатской работы наличие официально подтверждённого внедрения в крупной финансовой организации является сильным свидетельством востребованности тематики и наличием практической направленности результатов.

Результаты докладывались на пяти конференциях: двух всероссийских конференциях с международным участием и трех мероприятиях, заявленных как международные, включая SYRCoSE, «Радиоэлектронные устройства и системы для инфокоммуникационных технологий» и конференции по информационной безопасности и искусственному интеллекту.

5. Теоретическая и практическая значимость

Теоретическая значимость работы состоит в конкретизации нескольких способов применения машинного обучения к системным данным Linux в условиях ограниченных ресурсов. Для системных журналов предложена схема, в которой текстовое представление одиночного события используется как вход трансформерного классификатора, а детальная тактическая интерпретация выполняется только после первичного бинарного отбора. Для задачи паролей показано использование совокупности статистических признаков случайности как признакового пространства обучаемого классификатора. Отдельным важным результатом является организация потокового вывода, при которой вычислительная эффективность повышается не за счет изменения нейросетевой архитектуры, а за счет управления формированием и выравниванием пакетов входных последовательностей.

Практическая значимость заключается в разработке и внедрении на производстве программного комплекса, воплощающего предложенные методы.

6. Замечания к работе

Между тем работа не свободна от ряда недостатков

1. Ряд ключевых терминов используется без точных определений. В частности, не вполне ясно разграничены «потоковые события», «системные события» и «последовательности событий»; термины «контекстное представление», «семантическое упрощение» и «вычислительная сложность события», фактически задаются через конкретные приемы реализации, но не формализуются. Более строгая терминология сделала бы границы полученных результатов яснее.

2. Ряд аспектов требует дополнительных пояснений, например, «вредоносные полезные нагрузки, встроенные в обфусцированные команды» (стр. 12).
3. Требует пояснения оценка необходимой длины входной последовательности. В табл. 2.2 на с. 56 после предварительной обработки приведен 99-й процентиль длины 428,3 токена, тогда как далее, на с. 88–89, утверждается, что длины 128 токенов достаточно для покрытия более 99% данных. Возможно, здесь используются разные способы подсчета длины, однако из текста это не следует. Поскольку ограничение окна непосредственно влияет и на качество, и на производительность, данный момент необходимо объяснить.
4. Одна из основных гипотез работы – небольшие ресурсы для анализа системных событий – требует более тщательного обоснования.
5. Диссертационную работу, безусловно, усилило бы применение больших языковых моделей, даже в виде перспектив и предварительных оценок. Тем более, что последние модели (в частности, Cloude Mythos/Fable) демонстрируют многообещающие результаты в области безопасности.

Указанные замечания не снижают значимости полученных результатов и не влияют на общую положительную оценку диссертационного исследования.

7. Заключение

В целом, считаю, что диссертационная работа Русанова Михаила Андреевича соответствует требованиям пунктов 9-14 Положения о присуждении ученых степеней, утвержденного постановлением Правительства РФ от 24 сентября 2013 года № 842, предъявляемым к кандидатским диссертациям, а её автор заслуживает присуждения учёной степени кандидата технических наук по специальности 2.3.5 Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Профессор кафедры системного
программирования Санкт-Петербургского
государственного университета, доцент, доктор
технических наук

/Кознов Дмитрий Владимирович/

25.08.2026